The Merits of Multi-Hop Communication in Deep Space

Julian C. Breidenthal
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109-8099
818-354-7655
jbreid@jpl.nasa.gov

[1]*Abstract*— The physics of communication give an advantage to multi-hop systems, because the ratio of power received to power transmitted on a telecommunications link over a fixed distance S increases by a factor of n^2 if the link is divided in to n equal hops of length S/n. If terminal characteristics are left unchanged, the cost of a system with multiple hops is higher than for a single hop, but the resulting system has an efficiency-to-cost ratio n times higher. The greater efficiency can be applied to higher capacity, or reduced cost to obtain the same capacity. Multi-hop systems offer simple scalability strategies, capacity that scales proportional to the square of the investment, robustness to node failure, and the feasibility of obtaining multiple orders of magnitude increase of capacity in deep space using presently available technology. This conclusion remains true even when the additional noise introduced on each segment is taken into account, if modern error-control coding techniques are applied. Candidate mission designs based on a multi-hop approach are discussed for high-capacity Earth-Mars and interstellar communication systems.

TABLE OF CONTENTS

1. INTRODUCTION

A key challenge of the coming century is to develop a communication system between Earth and Mars with orders of magnitude greater capacity than present systems. The conventional approach to long-haul communication in deep space is to transmit signals in a single hop. The hop is bounded at the Earth end by a reflector antenna of diameter in the range 10 m to 100 m, a transmitter emitting power in the range 1 kW to 1 MW, and a receiver with an equivalent noise temperature[2] of 10 K to 100K. At the spacecraft terminal the antenna usually has a diameter in the range 0.1 m to 10 m, the transmitter emits a power in the range 10 mW to 100 W, while the receiver has a noise temperature of 100 K to 3,000K. There is also usually an error-detecting and correcting system on the hop from spacecraft to Earth while the Earth to spacecraft hop usually has only error detection. Typically the wavelength is in the range 0.5 cm to 20 cm.

The large difference in characteristics between the Earth and spacecraft terminals reflects the fact that it is much more expensive to put antenna area, transmitter power, low noise temperature, or decoders at a site in deep space than at a site on the Earth. The primary underlying factor is the cost per unit mass lifted from the Earth to deep space, although reliability is also a consideration. Under such a constraint, and assuming a single hop in the telecommunications link, a good strategy is to put all the mass and complexity on the Earth, while keeping the spacecraft as light and reliable as possible.

This may or may not be the optimum strategy. It is useful to ask, for the purpose of seeking an optimum, whether a system designed assuming a single hop results in the best deployment of telecommunications resources. Is it possible that a system designed with multiple hops would be superior? On the surface it would appear that additional terminals would cost more, decrease reliability, and introduce more noise. However, it will be shown that a multi-hop system can be a much more cost-effective system than one with a single hop, depending on the cost characteristics of a telecommunications terminal and the measures applied to handle reliability and noise.

2. ANALYSIS

Consider a single hop communications link over a distance S, as depicted in Figure 1. A power P_t is transmitted from an antenna of physical area A_t, and traverses the distance S. The power leaving the transmitting antenna can be characterized by an Effective Isotropic Radiated Power (EIRP), the power required from a fictitious transmitter radiating uniformly in all directions, which matches the power transmitted by the real transmitter-antenna combination in the direction of the receiver. However, the real antenna directs more of the power in some directions,

---

[1] 0-7803-5846-5/00/$10.00 © 2000 IEEE

[2] Equivalent noise temperature is a way of measuring the noise introduced by a receiving system. The noise power introduced depends on the width of the frequency band examined, and is calculated by the formula $P_n = kTB$, where P_n is the noise power, k is Boltzmann's constant 1.38×10^{-23} W/K-Hz, T is the equivalent absolute temperature in kelvins (K), and B is the bandwidth in Hz.

resulting in a gain G_t and $EIRP = P_t G_t$. The far-field gain of a well-pointed reflector antenna is [1]

$$G_t = \frac{4\pi\eta_t A_t}{\lambda^2} \quad \text{(dimensionless)} \tag{1}$$

where λ is the wavelength of the transmitted signal and η_t is the antenna efficiency (a fraction less than unity, typically in the range 0.4 to 0.8). This results in an illumination level W at the receiver

$$W = \frac{EIRP}{4\pi S^2} = \frac{P_t G_t}{4\pi S^2} \quad . \tag{2}$$

The receiving antenna intercepts P_r proportional to its area A_r and efficiency η_r. This is invariably something less than the full power transmitted, the remaining power passing around the receiving antenna and ultimately being lost to free space, wasted from a communications standpoint.

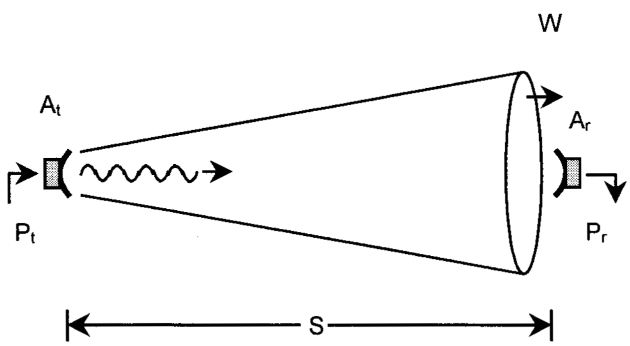

Figure 1. Single-hop Telecommunications Link.

One measure of the "efficiency" of the link is the ratio P_r/P_t. This represents the fraction of transmitted power available at the receiver to use for communication functions. Denoting this efficiency α and combining Equations 1 and 2 yields

$$\alpha = \frac{P_r}{P_t} = \frac{\eta_t \eta_r A_t A_r}{\lambda^2 S^2} \quad . \tag{3}$$

Now consider dividing the telecommunications link into n hops as shown in Figure 2. For the purpose of this paper, a hop carries data in one direction, and consists of the transmitting equipment at one location, the receiving equipment at another location, and the intervening space between them. The hops are distinguished from the terminals themselves, which except at the ends of the chain have both transmitting and receiving equipment at one location. The mid-chain terminal equipment is assumed to be arranged to forward data from the receiver to the transmitter in the direction of the next hop, and is assumed to possess error detection and correction capability so that errors introduced in reception are corrected before data is passed to the next hop.

For the moment the hops may be considered to be identical and unchanged from the original single hop except in the distance to be covered (other possibilities are discussed later). Observing that the distance to be covered is now shorter by a factor n, the new efficiency of each hop is

$$\alpha' = \frac{\eta_t \eta_r A_t A_r}{\lambda^2 (S/n)^2} = n^2 \alpha \quad . \tag{4}$$

Evidently efficiency improves by a factor of n^2.

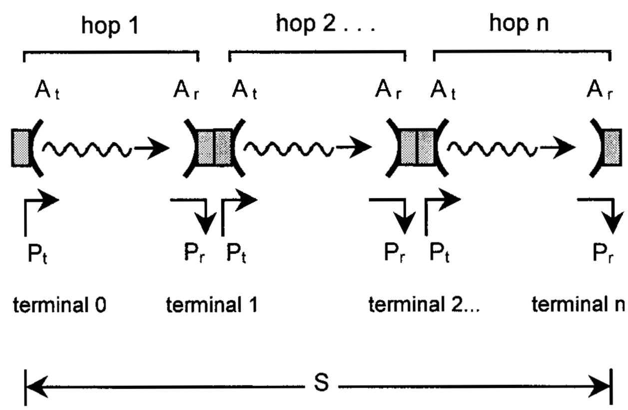

Figure 2. Multi-hop Telecommunications Link

However, efficiency is not the only concern. Cost matters as well. The costs will certainly depend on a large number of factors, technology being prominent among them, making it difficult to solve all situations in advance (this is especially true for as-yet-undiscovered technology). Despite this obstacle, a useful cost function to consider is one which takes into account the major features of the problem: the number of hops, a fixed cost u for each hop, and an adjustable, generalized, communications performance Q producing a variable cost per hop vQ^γ. The meaning of u and vQ^γ is intended to be adjusted from problem to problem depending on the significant factors involved. For instance, u could reflect the operating cost per hop, the minimum cost of putting a spacecraft at its station, basic cost of a transponder, or any other parts of the system which are considered fixed for the purpose of discussion. On the other hand, vQ^γ could reflect the cost of placing the mass associated with the parameters to be selected by a communications designer: antenna area, transmitter power, decoding capacity, receiver noise temperature, or wavelength. The exponent γ is included to allow for the possibility that cost per unit Q may increase or decrease as Q is increased. For example, larger antennas might be heavier and therefore more costly per unit area for structural reasons, or transmitters of greater output power might be more efficient and therefore less costly per unit output. The coefficient v merely converts between Q^{γ} and cost. Note that u and vQ^{γ} describe the cost of a hop, and therefore include the cost of part of one terminal and cost of part of another terminal.

It is important to realize that the precise details of the cost function are not important for remainder of this section. It is only important that the cost of a hop does not significantly increase as the number of hops is increased.

With these assumptions in mind, a cost function $Y(n,Q)$ is proposed:

$$Y(n,\vec{Q}) = \sum_{i=1}^{n} u_i + v_i Q_i^{\gamma} \quad . \tag{5}$$

Here the total cost Y is considered to be the dependent variable, n and the array of communications capacities Q_i are considered to be the independent variables, while the coefficients u_i , v_i , and γ are considered to be constants fixed by the problem domain.

Now consider the problem of assessing benefit and cost. Let the benefit-to-cost comparison be characterized by the ratio of efficiency to cost α / Y, and the relative merit M of one approach over another be defined by

$$M \equiv \frac{\alpha' Y}{Y' \alpha} \quad . \tag{6}$$

A value of M greater than unity indicates a profit from a benefit-cost standpoint; that is, the approach bearing the prime mark (') has a higher ratio of benefit to cost than the other approach.

Assuming, temporarily, that the hops in the divided link are identical and unchanged from the original single hop except in the distance to be covered, and denoting the divided link with the prime mark ('), the combination of Equations 3, 4, 5, and 6 indicate that the relative merit is

$$M = \frac{\alpha' Y}{Y' \alpha} = \frac{n^2 \alpha \left[u + vQ^{\gamma}\right]}{\alpha n \left[u + vQ^{\gamma}\right]} = n \quad . \tag{8}$$

Evidently a favorable change in the ratio of efficiency to cost, by a factor of n , occurs for dividing the link, despite the fact that the cost has increased by n . So not only is the divided link more efficient by a factor of n^2 , it is also more cost-effective by a factor of n .

3. Discussion

At this point one might reasonably argue that it is completely impractical to create hops in deep space that are "identical and unchanged from the original single hop except in the distance to be covered." Indeed, it is generally considered difficult to place a 34-meter diameter antenna (typical of the ground terminal) in deep space. However, a closer inspection of Equation 3 indicates that it is only the product of the transmitting and receiving areas which matters. Therefore one can achieve the same efficiency by downsizing the transmitting antenna, while upsizing the receiving antenna, in such a way as to keep the product of areas constant. As a special case consider the Mars Global Surveyor X-band (3.5 cm) telecommunications link (typical data rate 80 kbps), which has a pair of antenna diameters (d_t = 1.5 m , d_r = 34 m) at the spacecraft and Earth respectively. Assuming instead that equal diameters are desired, the same product of areas could be obtained with

$$d = \sqrt{d_t d_r} = 7.1 \quad \text{meters,} \tag{9}$$

which, while difficult, is more feasible than placing 34-meter antennas in deep space.

Note that it is not necessary to duplicate the efficiency of the original single hop in each of the multiple hops, because there is such a large benefit in both in efficiency and cost effectiveness with increasing n that efficiency losses could be made up by dividing the link into more hops. For concreteness, a set of n = 514 hops characterized by (d_t = 1.5 m, d_r = 1.5 m) would have the same cost-benefit ratio as a single hop characterized by (d_t = 1.5 m , d_r = 34 m), and it would have 514 times the efficiency. These smaller antennas would be even more feasible than those sized by Equation 9.

The possibility of increased efficiency at the same cost-benefit ratio is quite exciting from the viewpoint of communications capacity. Up to the present time, the data rate of deep space telecommunications links has been limited by the uncorrected error rate remaining after the best available error-control coding schemes have been applied. The uncorrected error rate is itself generally a function of the ratio of energy-per-bit E_b to noise spectral density N_0

$$E_b / N_0 = P_r / RkT = \alpha P_t / RkT \tag{10}$$

where P_r is the received (data) power[3], P_t is the transmitted (data) power, R is the data rate, k is Boltzmann's constant 1.38 x 10^{-23} W/K-Hz, and T is the equivalent absolute temperature of the receiver in kelvins (K). Therefore, all other things being equal, an increase in efficiency α allows a proportional increase in data rate R while maintaining the same error rate. In the n = 514 hop example above, the data

[3] The data power is usually a significant fraction of, but somewhat less than, the total power. This difference does not affect the scaling of data rate with efficiency, however, since both transmitted and received power obey the same ratio of data to total power.

rate could be increased 514 times, or from 80 kbps to 41 Mbps using Mars Global Surveyor as the reference.

There is admittedly an error-rate penalty associated with multiple hops. If left unaddressed, the error rate would increase proportional to the number of hops. One approach to overcoming this problem would be to increase n and thereby the efficiency, raising the signal-to-noise ratio. This is fruitful because codes are available [2] for which the error rate falls by two orders of magnitude with only a 0.2 dB (4.7%) increase in E_b / N_0 (see Figure 3). The number of hops would only have to be increased by about half this ratio (2.3%) to realize a two-orders-of-magnitude error reduction, because efficiency is proportional to n^2 . Alternatively the data rate could be dropped by 4.7%, or a re-transmission protocol could be used to correct for corrupted data. Thus there are at least three approaches to overcoming the error-rate penalty, and so to achieve nearly all of the benefit of a multi-hop system.

There is also a reliability penalty associated with multiple hops. All other things being equal, the failure rate increases proportional to the number of hops. However, the consequence of a failure is not as great as in a single-hop system, where loss of the hop means loss of the communication system. Instead, a multi-hop system offers the chance to continue operating at reduced data rate. If one terminal is lost, a double-length hop is created between the terminals adjacent to the failed one. This hop can operate at one-fourth the data rate of the undamaged system. If the terminals are maneuverable, over time they can be rearranged for uniform hop distance, raising the data rate to $[(n-1)/n\ ]^2$ of the undamaged system. Alternatively, the failed terminal could be replaced, restoring the full data rate. Thus, it is possible to heal the damage to a multi-hop system, and to continue operating during the healing process.

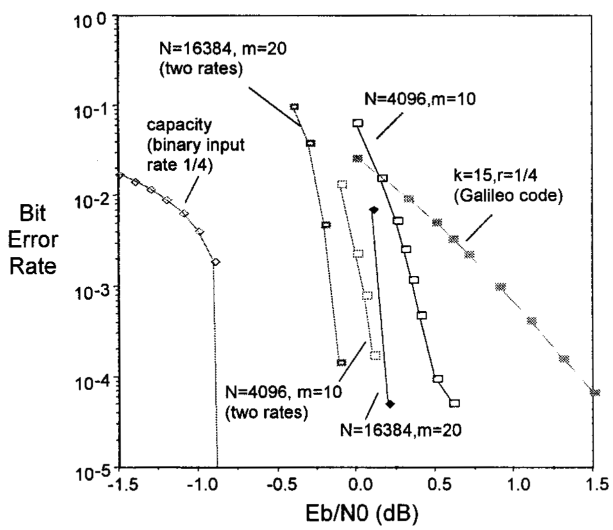

Figure 3. Bit-error Rate Performance of Some Turbo Codes. (Courtesy D. Divsalar [2])

Scalability of a multi-hop system can be achieved by starting with a few hops equally distributed along the communication path. The achievable data rate is necessarily low at this stage. The existing hops can then be divided by adding more terminals. The scaling should proceed by adding more hops, rather than by adding strings of hops in parallel, because the former case increases data rate by the square of the number of hops, while the latter case increases data rate only by the number of hops. That is, it is preferable to add extra terminals in series, rather than parallel. It will be seen later, however, that some installations suitable for the solar system necessarily include parallel strings.

For considerations of both reliability and scalability, it is helpful if the terminals have multiple data rates available, preferably continuously variable or at least a set matched to the likely values needed for healing or scaling.

Another exciting possibility is the application of the efficiency boost offered by multiple hops to the reduction of system cost. In this approach, the efficiency of the individual hops is intentionally downgraded to lower the cost of the hops, but only so far as can be managed while exceeding or matching the efficiency of a competing single-hop system. For an example, the antenna gain might be reduced to that of an isotropic radiator of effective area $\eta A = \lambda^2 / 4\pi$ (G = 1, close to a simple dipole), and n chosen to obtain the same efficiency as the Mars Global Surveyor example. This results in the same data rate, if powers, coding, and noise temperature are left unchanged. Using Equations 3 and 4, and setting the efficiencies equal, one obtains an equation to be solved for the value of n which would make up for the lower antenna gain,

$$\alpha' = \frac{\eta_t' \eta_r' A_t' A_r'}{\lambda^2 (S/n)^2} = \frac{\eta_t \eta_r A_t A_r}{\lambda^2 S^2} = \alpha \ . \qquad (11)$$

In this equation the prime mark (') denotes the divided link. Solving for n,

$$n = \frac{\sqrt{\eta_t \eta_r \pi^2 d_t d_r}}{4\lambda^2} = 56{,}500 \qquad (12)$$

for $\eta_t = \eta_r = 0.55$, $d_t = 1.5$ m , $d_r = 34$ m, and $\lambda = 3.5$ cm. To be competitive with the existing single-hop link with a cost around \$50M US, the terminals would have to cost less than approximately \$1K US. While this is a very low cost for any spacecraft, the low cost would be facilitated by the key features of the terminals: (1) absence of a need for attitude control or articulation, and (2) complexity comparable to a cellular telephone. With the addition of simple patch (G = 2) or horn (G = 10) antennas, the number of terminals and allowable cost per terminal would be (n = 14,100 ; Y = \$4K US) and (n = 565 ; Y = \$100K US), respectively. Of course, the higher gain antennas would place correspondingly greater demands for attitude control of the terminal. With G=2, attitude control to within approximately 90 degrees would be required, and with G=10 attitude control to within approximately 37 degrees would be required.

While these point solutions are interesting, it would be more valuable to seek the best number of terminals to achieve a fixed data rate. Here one considers the number of hops n to be an independent variable, then seeks to minimize the cost Y' for the divided link subject to constraints of data rate R, bit signal-to-noise ratio E_b/N_0, and distance S. Wavelength λ can be considered a constraint (as will be assumed in this paper), or with a slight modification of the equations below wavelength can be treated as one of the variables to be adjusted with generalized communication performance. Rearranging Equation 10 and applying Equation 3, one obtains

$$\frac{\lambda^2 S^2 R(E_b/N_0)}{n^2} = \eta_t \eta_r A_t A_r \frac{P_t}{kT} \equiv Q_i(n) = \frac{Q_1}{n^2} \quad . \tag{13}$$

The quantity $\eta_t \eta_r A_t A_r P_t / kT$ is identified with the hop's generalized communication performance Q_i from Equation 5, while the quantity $Q_1 = \lambda^2 S^2 R (E_b/N_0)$ is a constant of the problem which happens to be the value of the generalized communication performance required for a single hop. Assuming uniform hops and applying the cost model of Equation 5, the cost of the divided link is

$$Y'(n) = un + vnQ^\gamma(n) = un + vn\left(\frac{Q_1}{n^2}\right)^\gamma = u\left[n + \beta n^{(1-2\gamma)}\right] \tag{14}$$

where $\beta = Q_1^\gamma (v/u)$ is a scaled ratio of the coefficients of the variable and fixed costs per hop. The minimum of Y' with respect to n is sought, and this minimum is sought within a given context of technology with its own particular set of cost coefficients (u, v, γ). These coefficients are considered fixed in this optimization; that is, the optimization does not attempt to make comparisons between different technological contexts, only to find the best solution within a particular context.

Figure 4. System Cost as a Function of the Number of Hops.

An extremum of Y', if it exists, occurs at the number of hops n_0 such that

$$\left.\frac{dY'}{dn}\right|_{n_0} = u\left(1 + \beta(1-2\gamma)n^{-2\gamma}\right)\Big|_{n_0} = 0 \tag{15}$$

or

$$n_0 = [\beta(2\gamma - 1)]^{1/2\gamma} \quad . \tag{16}$$

The value of n_0 is real if $\gamma > 1/2$. The extremum is a minimum if

$$\begin{aligned}\left.\frac{d^2Y'}{dn^2}\right|_{n_0} &= u\left(\beta(1-2\gamma)(-2\gamma)n^{-(1+2\gamma)}\right)\Big|_{n_0} \\ &= u\left(2\gamma[\beta(2\gamma-1)]^{-1/2\gamma}\right) > 0\end{aligned} \tag{17}$$

which is true if $\gamma > 1/2$ (it is also assumed that $u > 0$ and $\beta > 0$).

This minimum is only beneficial if the cost is less than for a single hop, that is

$$Y'(n_0) < Y'(1) \tag{18}$$

or

$$u\left[n_0 + \beta n_0^{(1-2\gamma)}\right] < u[1+\beta] \quad . \tag{19}$$

Manipulation of Equation 19 leads to the constraint

$$\frac{(1+\beta)^{2\gamma}}{\beta} > (2\gamma-1)\left[1 + (2\gamma-1)^{-2\gamma}\right] \quad . \tag{20}$$

If this comparison is satisfied for a particular problem, then the divided link is less costly and the desirable number of hops is n_0 . Figure 4 can be used as an aid to visualizing the benefit when this constraint is satisfied, as well as the relationships among cost, β, and γ . As might be expected, the multi-hop approach has the best advantage in situations of large β (small fixed cost per hop compared to variable costs which rise with generalized hop performance) and in situations of large γ (cost of a hop rises steeply with generalized performance of the hop).

The foregoing optimization assumed a uniform cost of the hops. This is a good model for systems that are predominantly located on one planet surface, or predominantly located in a group of orbits which can be reached with little difference in launch energy. An interesting variation is one in which one hop has a drastically different cost, as in the hop between Earth (or Mars) and the space-based terminals. If one hop has a cost that is different by a multiplier p (p>0), that is, the hop costs $p(u+vQ^\gamma)$, then Equation 14 is modified to be

$$Y'(n) = u\left[(n-1+p) + \beta\left(n^{(1-2\gamma)} + (p-1)n^{-2\gamma}\right)\right] \quad . \tag{21}$$

This equation is more difficult to optimize analytically, but it can be optimized numerically. One sample of the resulting cost function is shown on Figure 4 marked with the value p=1/2 . As can be seen in the figure, there are still situations where a minimum of the cost occurs for n>1.

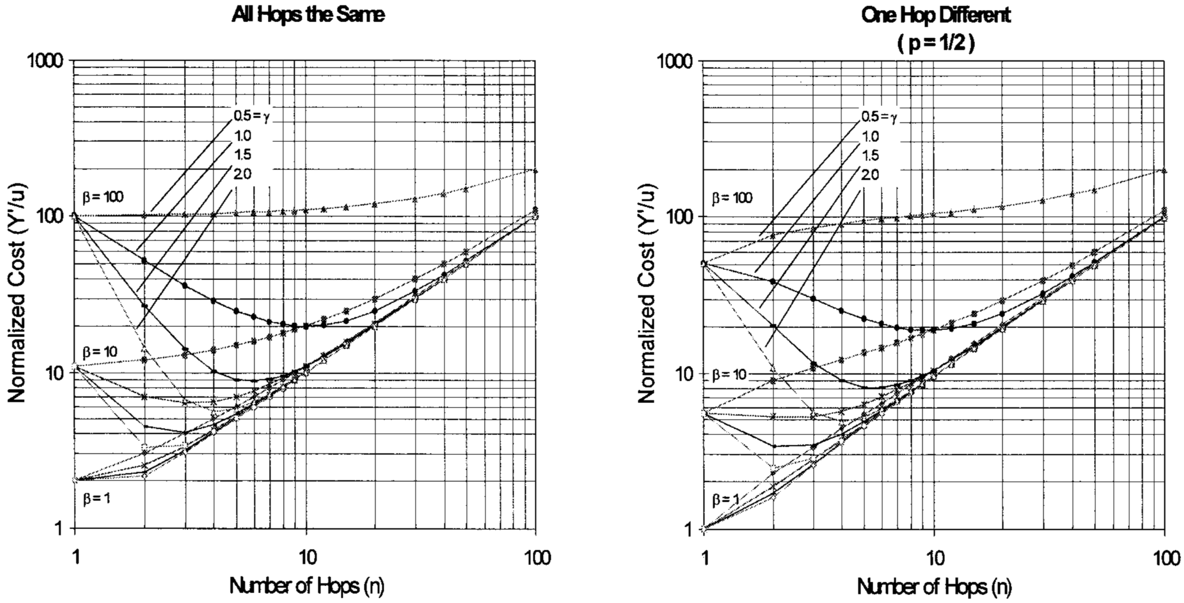

Figure 4. System Cost as a Function of the Number of Hops.

This value of p=1/2, while not unique, is significant in that it corresponds to situations where the cost of an earth terminal is zero, and the terminals are "symmetric" and "uniform". Here "symmetric" means that the terminals have equal fixed cost and equal contribution to generalized performance Q (and therefore variable cost) in the transmit and receive capabilities of a hop, and "uniform" means that all the terminals have the same communications system design, including the Earth terminal. This is a good approximation to situations for which the cost of putting capability on the Earth is nearly zero compared to putting it in space, and it is desired to apply Earth terminals of the type usually used for Earth-orbiting spacecraft to the deep space problem.

Although it is beyond the scope of this paper to examine specific systems, observe that it is likely that values $p<1/2$ are currently achieved in single-hop deep space systems. This probably occurs through the use of highly asymmetric designs emphasizing more contribution to capacity from the Earth terminal. If so, Equation 21 can justify selection of a single-hop approach when the lower ranges of β and γ prevail.

For real system tradeoffs, consideration should also be given to several other factors not explicitly considered here. These are:

- reduction of following mission costs due to smaller telecommunications equipment,
- reduction of fixed costs when non-recurring expenditures are amortized over terminals or missions,
- number of customer terminals needed, for visibility constraints or collection of data at the planet,
- number of Earth terminals needed, for visibility constraints or distribution of data directly to the customer, and
- enablement of missions whose payload would be prohibitive if a single-hop telecommunications system were included.

It is useful to note that an advantage of any multi-hop scheme in deep space is the absence of atmosphere effects on all but the planetary hops. This situation increases the utility of shorter-wavelength signals such as Ka-band or optical. Only the hops from planet to deep space (and in the case of Mars links, primarily the Earthbound terminal) would have to supply the extra performance to overcome atmospheric effects. As shorter wavelengths can yield more efficiency for the same antenna area (see Equation 3), they are popular for increasing system performance. Unfortunately, atmospheric factors have proved an obstacle to realizing all of the potential benefit for Earth-based terminals, which could be more easily overcome for a multi-hop system.

As a final point of discussion, note that the reasoning behind selecting (or not selecting) a multi-hop approach applies equally well to optical systems as to radio systems. The underlying reason for choosing a multi-hop system is still the same: the power density of a signal falls proportional to S^{-2} . In a detailed analysis, an adaptation would be needed in the formulas above to replace E_b / N_0 and kT with quantities appropriate for quantum-limited processes, such as bits-per-photon and background count rate. Additionally, the specific values of the cost coefficients (u, v, γ) will be different, as they might well be for radio systems at different wavelengths. Nonetheless, when the optical problem is reduced to a cost function such as Equation 14, the same character of possessing a term proportional to n^{-2} will appear because of the S^{-2} signal decay. The relative balance between fixed cost per hop and the n^{-2} term is what produces the opportunity for cost savings for any particular technology.

4. APPLICATIONS

A key challenge for a multi-hop communication system in deep space is the deployment of the terminals to locations suitable for communication, at reasonable cost. Within the gravitational influence of the sun, a terminal will be subjected to a steady gravitational force that would move the terminal away from any initial state of location and velocity. This force must be dealt with to maintain communications over an extended period of time. One strategy would be to apply a countervailing force by means of continuous thrusting or a tether, thereby to maintain the terminal in a fixed location relative to the ends of the link. However this approach does not appear feasible at this time.

A more economical approach is to accept the sun's gravitation as a given, and take advantage of the fact that the terminal will accelerating toward the sun at all times. Since the terminals will be moving, any particular terminal will at one time be nearer to one target of communication, while at another time it will be nearer to another target. The trajectory of the terminal need not be repeatable or even deterministic, although this paper will be restricted to considering constellations of terminals that are placed in stable orbits around the sun. The Earth-Mars communication problem will be taken as a special case to examine the concept, although there is no particular limitation against applying the techniques discussed to other planet pairs.

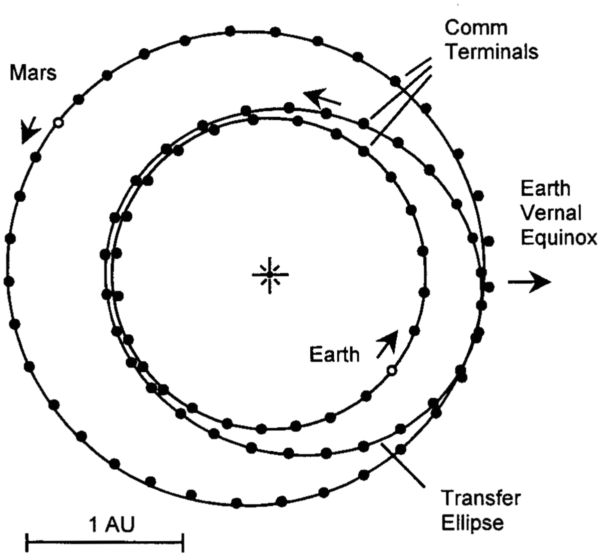

Figure 5. Elliptical transfer between planetary orbits.

One possible deployment of the terminals is to place them in three orbits about the sun: the orbit of the Earth, the orbit of Mars, and an elliptical orbit which contacts (or nearly contacts) the orbits of the Earth and Mars. Many terminals would be placed in each orbit, as depicted in Figure 5. The terminals in Earth and Mars orbit would have the same orbital period as their planet, and so maintain their orbit phase constant relative to their planet, moving approximately in lock-step around the orbit. There would be some compression of the distance between terminals near aphelion due to their smaller velocity there, and a corresponding expansion near perihelion, the effect being more pronounced for increasing eccentricity of the orbit. Communication proceeds from the Earth to the nearest terminal in the Earth's orbit, along the Earth's orbit to the nearest terminal in the "transfer ellipse," then along the transfer ellipse to the nearest terminal in Mars' orbit, then along Mars' orbit to the terminal nearest Mars and thence to Mars.

There are three disadvantages to the elliptical transfer between planetary orbits. Firstly, the communications path is far from a straight line. At worst case, with Mars at aphelion and Earth on the opposite side, the distance is increased from S_0 = 2.5 AU to $S = \pi(1 + 1.52 + 1.18) =$ 11.6 AU. This has the undesirable effect of lowering the efficiency by the same amount as if Mars were actually farther away. The efficiency decrease is by a factor $(S_0 / S)^2$ = 4.6 % compared to a direct route with the same number of hops.

Secondly, many of the terminals are in orbit away from the shortest path at any given time. By considering the worst-case communication distance, it can be seen that twice as many terminals must be installed to achieve a given spacing, doubling the cost. The terminals need not be wasted, however, since communication could be carried out along two paths, one shorter and one longer. Thus the communication capacity is doubled along with the cost, so the benefit-to-cost ratio remains the same. Nonetheless, one would prefer to have used those extra terminals instead to reduce spacing by a factor of two and thereby quadruple the capacity. Thirdly, a significant fraction of the terminals are redundant near to the Earth orbit. This could be overcome by using a second transfer ellipse as in Figure 6, instead of populating the Earth's orbit.

The elliptical transfer approach has the advantage that there are no particular limits to the density of terminals along the path, which makes it useful for large n. Another advantage is the fact that the transfer ellipse can conceivably be arranged to be the same as one frequently used for cruise between Earth and Mars. The terminals could then supply communications during the cruise phase of a mission just as they would at Mars, reducing the need for customer spacecraft to carry heavy communications equipment.

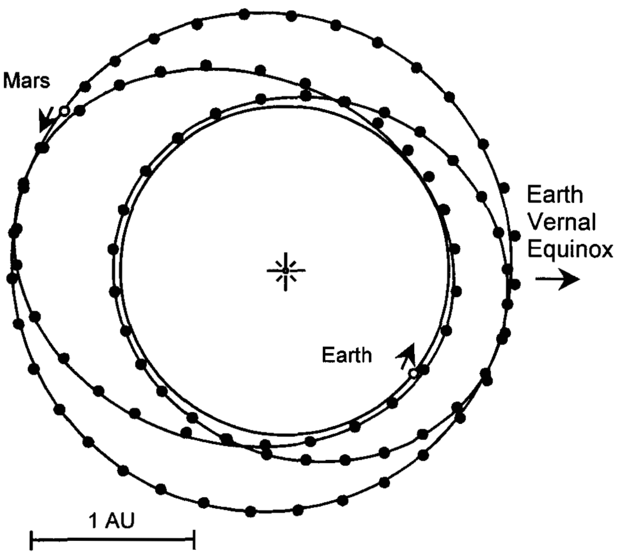

Figure 6. Two-Petal Elliptical Transfer Between Planetary Orbits.

A simple and elegant solution proposed by W. J. Hurd is the minimal Earth ring, depicted in Figure 7. Six terminals are deployed in the orbit of the Earth, spaced by one-seventh of a full circle (approximately 51.25 degrees) with Earth providing the seventh terminal. The distance between terminals is 0.87 AU and the maximum distance between Mars and a terminal is 0.87 AU with Mars at aphelion, centered between terminals. At worst case, with Mars at aphelion and Earth on the opposite side, the distance is increased from $S_0 = 2.5$ AU to $S = 3.5$ AU. The extra path length lowers the efficiency to $(S_0 / S)^2 = 52$ % compared to a direct route. Given that four hops are available along the maximum path, the efficiency would be increased by $(4S_0/S)^2$, for an eight-fold increase in data rate compared to a direct Earth-Mars link. This capacity would apply to both the long and short paths available through the ring, yielding a total capacity of 16 times a direct link.

It is desirable to find more efficient solutions than the elliptical transfer approach, and that still allow shorter hops than the minimal ring. In certain special cases there are solutions which are comparatively more efficient at particular terminal spacings. For instance, Figure 8 depicts a commutating ring, consisting of terminals placed on an intermediate orbit between Earth and Mars. The intermediate orbit bisects the distance between Earth and Mars at Mars aphelion and perihelion, has a semi-major axis of 1.26 AU and eccentricity 0.056, yielding a maximum distance of 0.33 AU from the orbit for either Earth or Mars. If such an orbit is populated with 24 terminals, the spacing between terminals will also be 0.33 AU and the maximum spacing to Mars will be about 0.37 AU when Mars is centered between two terminals. The greater length of the Mars hop could be made up by extra communication performance for the Martian terminal, or with a terminal following about one-half hop away in the orbit of Mars. Alternatively, with very little efficiency penalty, the number of terminals can be lowered to 21, increasing the maximum spacing to about 0.38 AU considering both Mars and the inter-terminal spacing. In this arrangement, the maximum path lowers the efficiency by a factor $(S_0 / S)^2 = (2.5\ AU / 4.69\ AU)^2 = 29$ % compared to a direct route. Given that 12-14 hops are available along the maximum path, the efficiency would be increased by $(12\ S_0 / S)^2$ to $(14\ S_0 / S)^2$, leaving the potential for a 40- to 56-fold increase in data rate compared to a direct Earth-Mars link. This capacity would apply to both the long and short paths available through the ring, yielding a total capacity of 80 to 112 times a direct link.

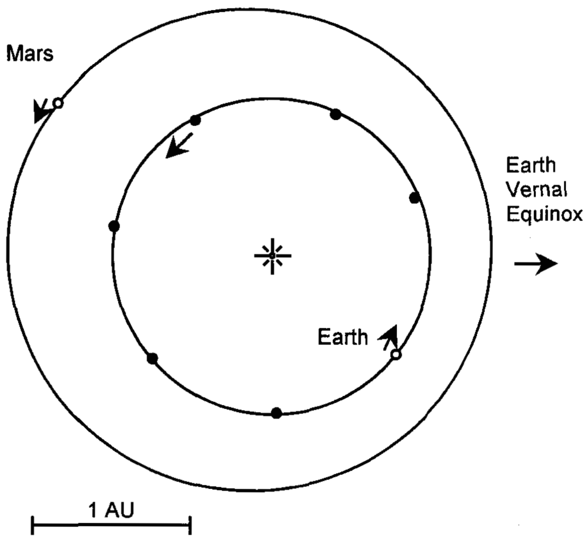

Figure 7. Minimal Earth Ring (courtesy W. J. Hurd).

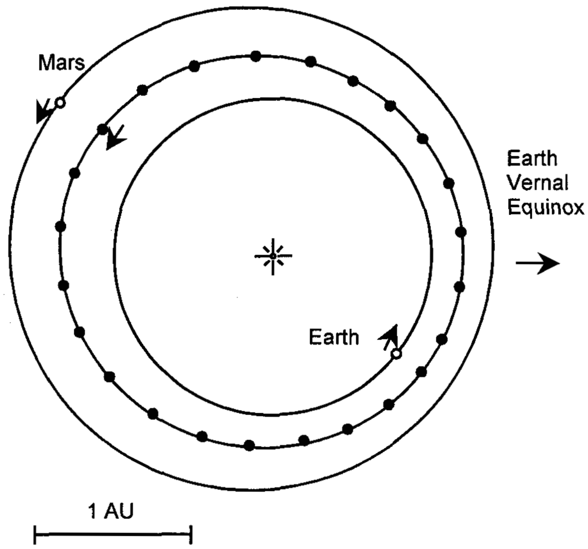

Figure 8. Commutating Ring.

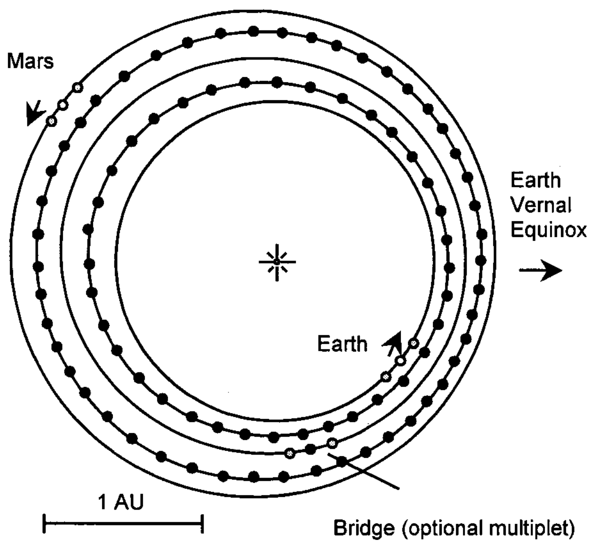

Figure 9. Two commutating rings with bridge.

The commutating ring concept can be extended to closer terminal spacing by placing more rings between Earth and Mars, as shown in Figure 9. Alternate rings can be populated by a single terminal, to serve as a "bridge" between the rings. The bridge (or the planets) can be augmented with a multiplet of nearby terminals as well, which serves to lower the maximum inter-terminal spacing and to carry extra capacity implied by the existence of long and short routes through any of the rings. The approach begins to lose value, however, as smaller terminals spacings are considered. This is because it becomes necessary to fill the entire area between Earth and Mars with terminals, most of which will not be participating in the two possible communications paths available at any given time on the innermost and outermost rings. At best, the extra terminals could be used as a weightless form of short-term data storage for bursts of data, routing the signals to take advantage of the significant time delay introduced by the speed of light. At worst, the extra terminals raise cost without adding to data rate.

The problem of decreasing the hop length in the distance between Mars (or Earth) and the basic commutating ring can be addressed more efficiently by placing a few terminals in "neighbor" orbits around Earth or Mars. For the purpose of this paper, a neighbor orbit is an orbit of a satellite about the sun with the same major axis as a reference body, but with a slightly different eccentricity and/or inclination. This situation is shown in Figure 10. By Kepler's law such an orbit has the same period as the reference body. In a rotating reference frame with a fixed line between the reference body and the sun, the separation between the reference body and the satellite will execute a repeating closed path at intervals of the orbital period, as depicted in Figure 11.

The perihelion of the neighbor orbit can be arranged to be between the reference body, e.g. Mars, and the commutating ring. Furthermore, multiple terminals can be placed in nearby neighbor orbits with phasing arranged to have their perihelia occur in a regular succession, so that there is always one between Mars and the commutating ring. With a single neighbor group each for Mars and the Earth to reach the same commutating ring described above, the maximum terminal spacing can be halved to 0.165 AU.

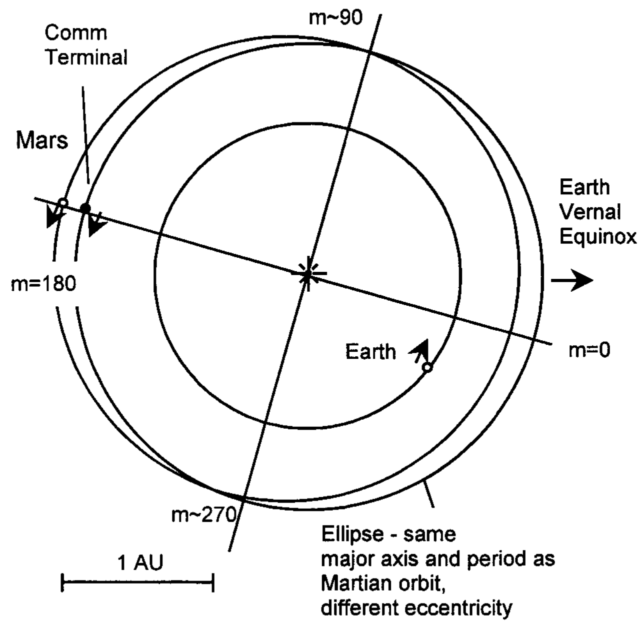

Figure 10. Neighbor Orbit -- Inertial View.

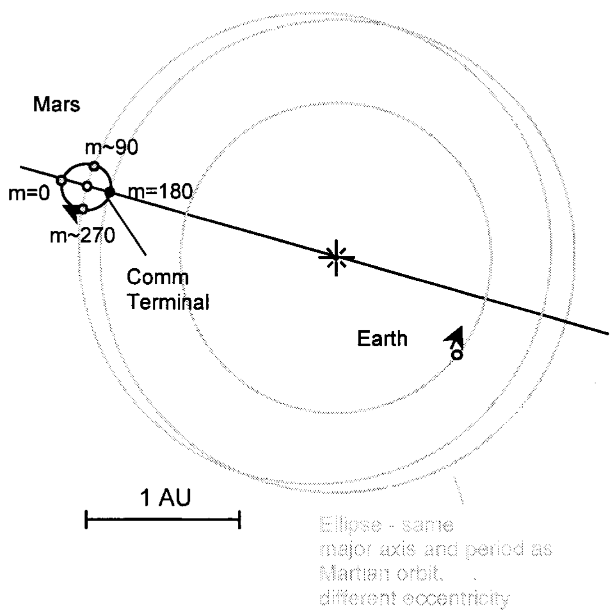

Figure 11 Neighbor Orbit -- Fixed Sun-Mars Line.

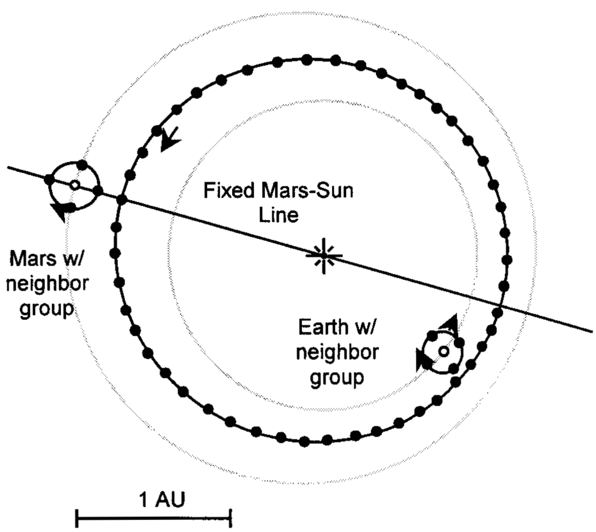

Figure 12. Commutating Ring with Neighbor Groups.

This situation is depicted in Figure 12, with four terminals each in the neighbor groups and 48 terminals in the commutating ring. In this arrangement, the maximum path lowers the efficiency by a factor $(S_0 / S)^2 = (2.5\ AU / 4.48\ AU)^2$ = 31 % compared to a direct route. Given that 28 hops are available along the maximum path, the efficiency would be increased by $(28\ S_0 / S)^2$, leaving the potential for a 244-fold increase in data rate compared to a direct Earth-Mars link. As in the previous cases, this capacity would apply to both the long and short paths available through the ring, yielding a total capacity of 488 times a direct link. If the terminals were sized with 7.1 m antennas as discussed for Equation 9 , and the remainder of the telecommunications system the same as Mars Global Surveyor X-band link, the capacity could be increased from 80 kbps to 39 Mbps. An additional bonus is that the efficiency to cost ratio would be 28 times better than the existing Mars Global Surveyor X-band link.

The neighbor approach offers the interesting side-benefit that a continuous view is available to one or more nearby terminals from all but the polar regions of the target planets. Furthermore, it may also be possible to reduce the terminal spacing further by using multiple neighbors of the Earth and Mars, neighbors of the satellites in neighbor orbits, or combinations of the two.

As a final case consider the problem of communicating to a nearby star, at say, 50 light years (3.1 million AU) distance. Solving Equation 10 for α and combining with Equation 3, one can show that the data rate R available on a single-hop link obeys

$$R = \frac{\eta^2 A^2 P_t}{kT(E_b/N_0)\lambda^2 S^2} \qquad (21)$$

assuming the transmitter and receiving antennas are the same. If one assumes the set of values ($\eta = 0.55$, $A = 1000\ m^2$, $P_t = 1$ kW, $T = 20$ K, $(E_b/N_0) = 1$ (0 dB) , $\lambda = 0.01$ m, $S = 3.1 \times 10^6$ AU $= 4.7 \times 10^{17}$ m), a data rate of 0.05 bps can be obtained. This set of values is selected to produce a bit rate which is at the lower limit of acceptability, where acceptability is defined as the onset of difficulty in maintaining bit synchronization due to local oscillator noise.

This set of values is such that the cost (and mass) associated with antennas, transmitters, and low noise amplifiers would far exceed that of the transponder or data handling equipment. If this is true, then β of Equation 14 would be much greater than unity *for a small number of hops* (this last qualification arises because terminal performance will fall with increasing number of hops, probably lowering β as the terminal cost declines toward the fixed cost of a transponder and data handling equipment). The value of β would be larger still if a choice were made for a 1 kbps or 1Mbps requirement, because the antenna area and/or transmitter power would be larger. Furthermore it is likely that γ is greater than unity, considering challenges posed by antenna stiffness and attitude control, transmitter waveguides and waste-power radiators, and a chiller for the low-noise radio components. If so, then it can be seen in Figure 4 or Equation 20 that a minimum cost system would not occur for a single hop, but for multiple hops instead. Note that this argument applies to radio only; no claims are made about the likely situation for optical communication.

The mission design for an interstellar multi-hop system is straightforward; the terminals can be placed along the hyperbola leading from the sun to the target. Since the total distance is large compared to the influence of the sun's gravitation, most of distance will be spanned by the part of the hyperbola that is very close to its asymptote. Most of the terminals launched along this path will very nearly be in a straight line and therefore provide nearly an ideal efficiency. The resulting "string of pearls" is depicted in Figure 13. The deployment into the string could proceed in one of three ways:

- by continuous launches as the population moves away uniformly, never to be decelerated,

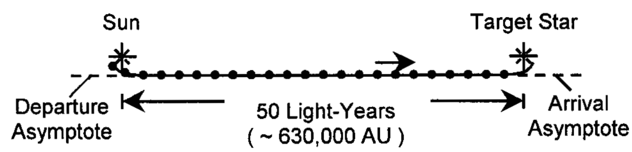

Figure 13. String of Pearls.

- by a single launch of all terminals at different velocities (slower velocities for terminals destined for nearness, faster for those destined far away), never to be decelerated, or

- by launching and decelerating the terminals to provide approximately fixed positions along the hyperbola.

The choice of which method would depend on the expected duration of service. Since deceleration is as expensive as acceleration, the first two options would be preferrable for short-duration missions.

5. CONCLUSIONS

A multi-hop approach to deep space communication offers the possibility to obtain data rate performance that grows as the *square* of the number of hops, and an efficiency-to-cost ratio that grows proportional to the number of hops. The benefits are applicable regardless of the technology, as long as signal power decays with the square of distance. Additional noise introduced on the hops can be overcome if modern error-control coding is applied with little loss of efficiency. A multi-hop system is more robust to terminal failure than a single-hop system, because a multi-hop system degrades in performance as terminals are lost rather than going out of action completely. If mechanisms for managing reliability are available, such as maneuverable terminals, replacement terminals, or variable data rates, the loss of data rate capability is limited to $[(n-1)/n]^2$ compared to the undamaged system. Multi-hop systems can be scaled, even after initial installation, by adding more terminals, preferably in series rather than in parallel. A generalized cost function is proposed which takes into account the major features of the problem: the number of hops, a fixed cost u for each hop, and an adjustable, generalized, communications performance Q producing a variable cost per hop vQ^{γ}. Examination of this cost function demonstrates that the efficiency boost offered by multiple hops can yield a lower overall system cost for a given data rate, if (a) the fixed cost per hop is small compared to the cost of antennas, transmitter power, or low noise temperature, and (b) the cost of generalized communications performance grows faster than $Q^{1/2}$. This conclusion applies equally well to radio or optical systems, since both possess the same signal decay with distance.

Mission designs are proposed which could apply the multi-hop approach to the interplanetary communication challenge. The designs presented here were limited to terminals placed in stable orbits about the sun, consisting of the planet orbits, elliptical transfer orbits, commutating rings between the target bodies, and neighbor orbits of various bodies. In particular, a single commutating ring between the Earth and Mars populated by 48 terminals, augmented by four terminals each in neighbor orbits of Earth and Mars is interesting. This arrangement could increase the data rate from the existing 80 kbps to 39 Mbps using roughly the same technology as was used for the Mars Global Surveyor X-band link. An additional bonus is that the efficiency to cost ratio would be 28 times better than the existing link.

Application of the analysis to the interstellar radio communication problem indicates that it is likely that a multi-hop system would be less costly than a single-hop system.

REFERENCES

[1] Gary Gordon and Walter L. Morgan, *Principles of Communications Satellites*, pp. 35-39, New York: Wiley, 1993.

[2] D. Divsalar and F. Pollara, "Turbo Codes for Deep-Space Communications," *TDA Progress Report 42-120, 29-39*, February 1995, Jet Propulsion Laboratory.

ACKNOWLEDGEMENT

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

The author wishes to thank C. D. Edwards, R. L. Horttor, W. J. Hurd, P. W. Kinman, A. Makovsky, C. T. Stelzried, and A. A. Wolf for many helpful comments and suggestions during preparation of the paper.

Julian (Jay) C. Breidenthal is a senior system engineer with the Telecommunications and Mission Operations Directorate of the Jet Propulsion Laboratory. He is presently working on the requirements and design of the Deep Space Mission System, popularly known as the "InterPlanNet", and has served as engineer, lecturer, and consultant in the areas of radio navigation, radio science, radio interferometry, electromagnetics, medical and power electronics, communications, monitor and control, mission operations, software engineering, and requirements management. Jay obtained his B.A. in Physics from U.C. Berkeley and an M.S. in Physics from MIT. During his (copious) free time he enjoys church activities, music, cycling, and hiking.